Microsatellite markers for *Eucalyptus pilularis* (Subgenus *Eucalyptus*); sourcing genetic markers outside the subgenus

By M. Shepherd[1),4),*)], C. Bihua[2),4)] and R. Henry[3),4)]

(Received 5th March 2013)

Abstract

Microsatellite markers remain the most broadly used molecular marker in eucalypt genetics. A major advantage of microsatellite markers is that they often transfer readily between related taxa circumventing the need to develop new markers *de novo* in each species. Markers have been developed for a number of species of major economic importance, mainly from the Subgenus *Symphyomyrtus*, but these may also be available for use in species of lesser economic importance from other subgenera. Here we report on the sourcing of microsatellite markers for *E. pilularis* (Subgenus *Eucalyptus* (Formerly *Monocalyptus*)) from species outside the subgenus. Ninety-seven precent (60 out of 62) of loci that amplified in the source taxon (*E. grandis*) also amplified in the target taxon *E. pilularis*. By characterising them on a diversity panel (n = 24) and a pedigree, a subset of 41 loci were distilled out that could be scored reliably and were polymorphic (Mean unbiased heterozygosity = 0.81). Predictions of efficient microsatellite marker transfer among eucalypts based on low evolutionary divergence have largely been borne out and are congruent with accumulating evidence of low sequence divergence within *Eucalyptus*. Upon this favourable background for microsatellite marker transfer, this study indicates highly efficient transfer is possible by identifying loci with broad PCR optima and adoption of approaches that favour cross-species transfer.

Key words: Monocalyptus, *Eucalyptus*, *Symphyomyrtus*, genetic diversity, microsatellite marker, internal transcribed spacer, ITS CCR,

Introduction

Microsatellite markers remain the most broadly used molecular markers in eucalypt genetics (Grattapaglia *et al.,* 2012). Their adoption as the preferred marker type for linkage mapping, studies of population diversity and structuring, and relationships amongst individuals and populations of eucalypts, paralleled the uptake in many other species of plants, animals and fungi (Barbara *et al.,* 2007; Grattapaglia and Kirst, 2008; Schlotterer, 2001). A major advantage of microsatellite markers is that they often transfer readily between related taxa circumventing the need to develop new markers *de novo* in each species, hence they may be available for use in studies of species of less economic importance (Peakall *et al.,* 1998). Forest tree geneticists have benefited particularly from this feature as it has been widely observed that transfer is most successful for species with long generation times, with mixed, or outcrossing breeding systems (Barbara *et al.,* 2007), and early predictions of high transfer rates among eucalypts has largely been borne out, and no doubt promoted their widespread application across the group (Byrne *et al.,* 1996; Jones *et al.,* 2001; Kirst *et al.,* 1997).

Eucalyptus pilularis Smith (Blackbutt) is a common, tall forest tree found in a wide latitudinal band (25–37 °S) on the east coast of Australia. It is the principal hardwood species for sawn timber in northern NSW (Henson et al., 2007) and along with *E. pyrocarpa*, it constitutes the Section *Pseudophloius*, Subgenus *Eucalyptus*, in the most recent classification of eucalypts by Brooker (2000). We have recently investigated the geographical and historical determinants of genetic structure in *E. pilularis* (Shepherd et al., 2010), and the degree of its relationship with *E. pyrocarpa* (Shepherd and Raymond, 2010) based on genetic variation at microsatellite loci.

Here we report on the sourcing of microsatellite markers for *E. pilularis* from outside the *Eucalyptus* subgenus (Formerly *Monocalyptus*). The eucalypts, in the broad sense, are a group of some 900 species classified into three genera, *Eucalyptus*, *Corymbia* and *Angophora* (Hill and Johnson, 1995), with most species of economic importance belonging to either the subgenus *Eucalyptus* or *Symphyomyrtus* (Eldridge *et al.,* 1994). The few members of the *Eucalyptus* subgenus of economic importance, however, means that most genomic resources have been developed in species from subgenus Symphyomyrtus (e.g. *E. grandis* and *E. globulus*) (Poke *et al.,* 2005). For example, a recent review of available genomic resources for the Myrtaceae showed that large numbers (i.e. > 50) of microsatellite markers have been developed in *Symphyomyrtus* species only (Grattapaglia *et al.,* 2012).

In one of the largest studies of trans-subgeneric transfer in eucalypts, we test the transfer of 71 loci sourced from *E. grandis* or *E. urophylla* (Brondani *et al.,* 2002; Brondani *et al.,* 1998; Brondani *et al.,* 2006) to *E. pilularis* and report genetic diversity parameters for one SSR marker developed *de novo*. We found comparatively high rates of inter-subgeneric transfer (97 %) compared to most previous studies in eucalypts (Byrne *et al.,* 1996; Glaubitz *et al.,* 2001; Jones *et al.,* 2001; Kirst *et al.,* 1997; Shepherd *et al.,* 2006; Steane *et al.,* 2001).

[1)] Southern Cross Plant Science, Southern Cross University, PO Box 157 Lismore, NSW 2480, Australia.
[2)] Fujian Academy of Forestry Sciences, Fuzhou, Fujian 350012, China.
[3)] Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, St Lucia 4072 Australia.
[4)] Former address: Centre for Plant Conservation Genetics, Southern Cross University, PO Box 157 Lismore, NSW 2480, Australia.
[*)] Corresponding Author: Dr. Mervyn Shepherd. E-Mail: Mervyn.Shepherd@scu.edu.au.

This we attributed to a filtering of the tested loci for a subpopulation of PCR robust loci and adoption of reaction conditions favouring transfer in our study.

Methods

Study design and Materials

A two stage assessment was used to test for microsatellite marker transfer to *E. pilularis*. In Stage 1, a set of 71 markers identified in *E. grandis* or *E. urophylla* (*Table 1*) and a single *E. pilularis* marker were screened on a set of six diverse *E. pilularis* individuals along with two *E. grandis* individuals to test compatibility of the loci with standardised PCR conditions (confirmed by amplification of the positive control *E. grandis*) and conservation of priming sites (confirmed by amplification of *E. pilularis*).

In Stage 2, a more rigorous assessment of the 24 most promising loci was conducted by screening an expanded set of 16 diverse *E. pilularis* individuals plus the two parents and six F1 individuals from a controlled pollinated cross (samples provided by Forests NSW) (total of 24 individuals; *Table 2*). The *E. pilularis* selected for this study were chosen to represent a diverse set of loca-

Table 1. – List of the 72 microsatellite markers used in the study. The flurophore and PCR program is listed with each marker as well as the amplification success of the marker (rating), degree of polymorphism and allele size range based on amplification in six *E. pilularis* individuals.

Index	*Marker*	*Fluorophore*	*PCR prog*[1]	*Marker rating*[2]	*Marker poly.*[3]	*UH_e*[4]	*Allele Range (bp)*
1	EMBRA 075	PET	PtTX58	0	-	-	
2	EMBRA 107	PET	PtTX54	0	-	-	
3	EMBRA 069	VIC	PtTX60	0	-	-	
4	EMBRA 080	VIC	PtTX58	0	-	-	
5	EMBRA 082	VIC	PtTX56	0	-	-	
6	EMBRA 085	VIC	PtTX60	0	-	-	
7	EMBRA 089	VIC	PtTX56	0	-	-	
8	EMBRA 230	Ned	PtTX50	0	-	-	
9	EMBRA 055	PET	PtTX54	0	-	-	
10	EMBRA 002	Ned	PtTX58	1	-	-	
11	EMBRA 120	PET	PtTX56	1	-	-	
12	EMBRA 020	Ned	PtTX56	2	M	-	147
13	EMBRA 058	6Fam	PtTX58	2	M	-	136
14	EMBRA 068	6Fam	PtTX54	2	M	-	106
15	EMBRA 092	6Fam	PtTX56	3	P	-	
16	EMBRA 113	Ned	PtTX56	3	P	-	
17	EMBRA 057	VIC	PtTX50	3	P	-	
18	EMBRA 095	6Fam	PtTX56	3	P	-	
19	EMBRA 224	Ned	PtTX60	3	P	-	
20	EMBRA 054	VIC	PtTX54	3	P	-	
21	EMBRA 172	PET	PtTX60	3	P	-	
22	EMBRA 207	PET	PtTX60	3	P	-	
23	EMBRA 061	VIC	PtTX50	3	P	-	
24	EMBRA 183	6Fam	PtTX54	3	P	-	
25	EMBRA 003	Ned	PtTX58	3	P	-	
26	EMBRA 150	Ned	PtTX50	3	P	-	
27	EMBRA 167	PET	PtTX50	3	P	-	
28	EMBRA 014	VIC	PtTX58	3	P	-	
29	EMBRA 034	VIC	PtTX54	3	P	-	
30	EMBRA 062	6Fam	PtTX50	3	P	-	
31	EMBRA 166	Ned	PtTX50	4	P	0.91	83-108
32	EMBRA 081	PET	PtTX54	4	P	0.68	80-94
33	EMBRA 001	VIC	PtTX56	4	P	0.91	104-136
34	EMBRA 040	6Fam	PtTX56	4	P	0.92	106-169
35	EMBRA 116	6Fam	PtTX50	4	P	0.47	120-122
36	EMBRA 066	6Fam	PtTX50	4	P	0.95	131-172
37	EMBRA 225	6Fam	PtTX56	4	P	0.95	155-185
38	EMBRA 094	PET	PtTX54	4	P	0.80	220-221
39	EMBRA 048	PET	PtTX58	4	P	0.76	91-111
40	EMBRA 043	VIC	PtTX56	4	P	0.94	97-124
41	EMBRA 011	Ned	PtTX56	5	P	0.97	126-159
42	EMBRA 007	6Fam	PtTX56	5	P	0.17	124-126

Table 1. – Continued.

Index	*Marker*	*Fluorophore*	*PCR prog*[1]	*Marker rating*[2]	*Marker poly.*[3]	UH_e[4]	*Allele Range (bp)*
43	EMBRA 083	6Fam	PtTX56	5	P	0.32	68-72
44	EMBRA 053	6Fam	PtTX56	5	P	0.94	121-134
45	EMBRA 103	Ned	PtTX54	5	P	0.80	189-207
46	EMBRA 039	VIC	PtTX56	5	P	0.38	120-122
47	EMBRA 028	6Fam	PtTX56	5	P	0.96	131-220
48	EMBRA 008	Ned	PtTX56	5	P	0.88	143-169
49	EMBRA 187	6Fam	PtTX58	5	P	0.85	183-189
50	EMBRA 012	6Fam	PtTX60	5	P	0.92	109-148
51	EMBRA 941	Ned	PtTX60	5	P	0.91	228-244
52	EMBRA 088	Ned	PtTX56	5	P	0.89	107-135
53	EMBRA 195	Ned	PtTX50	5	P	0.65	192-203
54	EMBRA 2000	PET	PtTX60	5	P	0.53	237-239
55	EMBRA 2002	PET	PtTX60	5	P	0.30	243-252
56	EMBRA 206	PET	PtTX58	5	P	0.77	273-302
57	EMBRA 098	PET	PtTX54	6	P	0.96	201-230
58	EMBRA 204	PET	PtTX56	6	P	0.89	136-156
59	EMBRA 042	PET	PtTX56	6	P	0.94	107-141
60	EMBRA 145	6Fam	PtTX60	6	P	0.17	107-109
61	EMBRA 175	6Fam	PtTX56	6	P	0.86	58-83
62	EMBRA 210	PET	PtTX50	6	P	0.89	205-223
63	EMBRA 182	VIC	PtTX50	6	P	0.74	170-222
64	EMBRA 214	Ned	PtTX54	6	P	0.53	114-130
65	EMBRA 169	PET	PtTX54	6	P	0.84	117-143
66	EMBRA 070	6Fam	PtTX56	6	P	0.95	114-151
67	EMBRA 209	6Fam	PtTX58	6	P	0.53	115-132
68	EMBRA 164	Ned	PtTX60	6	P	0.95	113-144
69	EMBRA 006	VIC	PtTX58	6	P	0.53	121-135
70	EMBRA 104	PET	PtTX58	6	P	0.88	125-146
71	EMBRA 242	PET	PtTX56	6	P	0.69	143-150
72	EPILMYB2	6Fam	PtTX60	6	P	0.91	150-175
Mean (SD) for 41 transferred and scorable loci (i.e. excl. EPILMYB2)						0.75(0.23)	

[1] PCR Program See AUCKLAND *et al.* (2002) for explanation of codes.
[2] Marker rating. A rating from 0 to 6 where 6 was optimal. See Methods.
[3] Marker Polymorphism in *E. pilularis*. - = no amplification; M = Monomorphic; P = Polymorphic.
[4] Unbiased Expected Heterozygosity = Nei's unbiased heterozygosity (NEI, 1978) for small sample sizes estimated in POPGENE v 1.31.

tions and taken from a larger range-wide collection of *E. pilularis* made for a study of population diversity (SHEPHERD *et al.,* 2010).

At both Stages, loci were assessed for polymorphism and assigned a marker rating between 0 and 6 based on whether they transferred to the target taxon (*E. pilularis*), their degree of polymorphism, and amenability to automated scoring with GeneMapper V3.7 software (Applied Biosystems). Marker ratings were; 0 = no amplification in source taxa, 1 = amplification in source taxon only, 2 = amplification in target taxon but monomorphic, 3 = amplification in target taxon, polymorphic but can't score, 4 = amplification in target taxon, polymorphic, manual score, 5 = amplification in target taxon, polymorphic, assisted-automated score, 6 = amplification in target taxon, polymorphic, automated score. Population diversity parameters were estimated using GenAlEx v6 (PEAKALL and SMOUSE, 2006).

At Stage 2, additional checks for Mendelian inheritance and the presence of null alleles were carried out using data from the F1 progeny. Genotypes were checked for congruence with their parental genotypes with each progeny expected to contain one allele from each parent. The only inconsistencies evident could be explained by the presence of a null allele in one of the parents. Here an offspring appeared as an apparent homozygote where it was expected to be heterozygous based on parental genotypes.

DNA extraction

DNA was extracted from frozen or air dried foliage using the DNeasy 96 plant kit (Qiagen Pty Ltd, Doncaster, VIC). The frozen tissue protocol was used with the following modifications; 2% PVP-40 and 2.5 mg/ml charcoal were added to 600 µl of Buffer API. The cell lysis step 65 °C was increased to 30 min and an additional wash step with 100% ethanol was included. DNA was eluted in a final volume of 100 µl of AE buffer (10 mM Tris-Cl, 0.5 mM EDTA; pH 9.0). Starting material consisted of 8 leaf disks (~50 mg of frozen or 10 mg of dried tissue) punched from the leaf using a hole punch. DNA was quantified using Picogreen dye by comparing fluorescence of samples with standard amounts of dsDNA with fluorescence measured using an MWGt Sirius HT-TRF Fluorometer/spectrophotometer (MWG, Ebersberg, Germany). Total yields were 800 ± 500 ng (mean ± SD).

Microsatellite markers and PCR conditions

A total of 71 microsatellite loci from the target taxa (*E. grandis* or *E. urophylla*) were tested for transfer to *E. pilularis* (*Supplemental material 1*). Additionally, the successful amplification of a microsatellite marker named EPILMYB2, (developed from a clone of the 2 gene from *E. pilularis* was also tested. The forward and reverse primers sequences were 5'-ATTTGCCCTCCTTCCCTTT-3' and 5'-CAGAATAATGAAAGAACAACACACA-3', respectively.

Microsatellite loci were amplified using a standardised PCR protocol described previously (SHEPHERD *et al.*, 2006) with the PCR buffer modified to include bovine serum albumin (BSA) at the rate of 0.6 ug/µl and the amount of template reduced to 1 ng per reaction in a total reaction volume of 12 µl. For Stages 1 and 2 of the transfer study (see below), each locus was amplified

Table 2. – Listing of 24 *E. pilularis* individuals used for the microsatellite transfer study. Sixteen individuals were selected from diverse locations (listed in increasing latitude) and eight individuals represent a pedigree of two parents and six controlled pollinated F_1 offspring.

ID	*Genotype*	*Provenance (State, collection location or nearby township*	*Lat (dec)*	*Long (dec)*
10694	k6737	QLD Nth. Fraser Island	-25.22	153.23
10690	k6733	QLD Sth. Fraser Island	-25.51	153.12
10672	k6715	QLD Cooloola	-25.96	153.11
4603	clyb bso 340	QLD Benarkin/Taromeo	-26.89	152.14
4613	clyb bso 350	QLD Googa	-26.97	152.01
4618	clyb bso 355	QLD Deongwar	-27.28	152.24
11019	ddpil9	NSW Double Duke	-29.17	153.21
10982	newpil8	NSW Newfoundland	-29.93	153.10
10965	conpil5	NSW Conglomerate	-30.12	153.05
10995	bbpil5	NSW Broken Bago	-31.51	152.70
10932	bby002	NSW Yerriyong SF	-35.02	150.58
10713	bbc003	NSW Currawon	-35.60	150.18
10731	bbm010	NSW Moruya	-35.79	150.07
10740	bbb009	NSW Bodalla	-36.15	150.07
10749	en008	NSW Nullica	-36.99	149.90
10703	ep003	NSW Eden	-37.01	149.91
10409	10135	Controlled cross parent 1	na	na
10410	20047	Controlled cross parent 2	na	na
10411	pol/k15	CC 10135 x 20047 - E.PIL	na	na
10412	pol/k17	CC 10135 x 20047 - E.PIL	na	na
10413	pol/k18	CC 10135 x 20047 - E.PIL	na	na
10414	pol/k19	CC 10135 x 20047 - E.PIL	na	na
10415	pol/k21	CC 10135 x 20047 - E.PIL	na	na
10416	pol/k22	CC 10135 x 20047 - E.PIL	na	na

individually then pooled for analysis on an Applied Biosystems 3730 Genetic Analyser (Southern Cross Plant Genomics, Southern Cross University, Lismore, Australia). The forward primers of each primer-pair were labelled with one of the Applied Biosystems G5 dye set. Raw electropherogram data was processed using GeneMapper V3.7 software (Applied Biosystems) and exported to an Excel spreadsheet (Microsoft).

Table 3. – Amplification conditions and genetic diversity for 24 loci based on a set of 16 *E. pilularis*.

Marker	T_a[2] (C)	Min allele size (bp)	Max allele size (bp)	Regular allele interval[3]	Marker Rating[5]	UHe	Multiplex
EMBRA 006	56	121	125	Y	6	0.33	B
EMBRA 008	56	128	178	N	5	0.92	B
EMBRA 011	56	120	163	N	6	0.92	C
EMBRA 012	58	108	148	N	6	0.90	A
EMBRA 028	56	56	86	N	6	0.84	C
EMBRA 039	56	114	123	Y	4	0.18	
EMBRA 042	56	117	162	N	5	0.94	C
EMBRA 053	56	109	143	N	5	0.88	C
EMBRA 070	56	114	181	N	4	0.87	
EMBRA 098	54	199	254	N	4	0.95	
EMBRA 103	54	192	207	N	4	0.89	
EMBRA 104	56	120	154	N	6	0.78	B
EMBRA 164	58	113	148	N	6	0.93	A
EMBRA 169	54	113	158	N	5	0.93	D
EMBRA 175	54	57	87	N	6	0.83	D
EMBRA 182	50	156	223	Y (M)	3	na	
EMBRA 187	58	181	215	Y	6	0.82	A
EMBRA 204	56	na	na	na	1	na	
EMBRA 209	56	115	132	Y	6	0.63	B
EMBRA 210	54	186	233	N	6	0.95	D
EMBRA 214	54	114	130	Y	6	0.63	D
EMBRA 242	56	140	151	N	4	0.78	
EMBRA 941	60	220	244	Y	4	0.88	
EPILMYB2[1]	58	150	177	N	6	0.73	A
Mean						0.80	
SE						0.04	

[1] Locus identified in *E. pilularis*. See methods.
[2] T_a – Touchdown program starting annealing temperature.
[3] Regular allele intervals were recorded as "yes" if allele sizes fitted the expected size according to the size of the repeat motif i.e. 2 bp pair intervals for a di-nucleotide repeat motif. An (M) indicates possible multilocus primer-pair.
[4] UH_e is Nei's unbiased expected heterozygosity for a small sample (NEI, 1978),
[5] Marker rating. Each marker was assigned a rating of 0–6 with 6 the most optimal based on transfer and scoring attributes. See Methods.

An independent estimate of the divergence amongst eucalypt groups using a fragment of the CCR gene.

A total of 55 CCR sequences from *Eucalyptus* (n = 35), *Corymbia* (n = 18) and *Angophora* (n = 2) were aligned manually using the alignment explorer of Mega v4 (TAMURA et al., 2007). Fifty sequences were obtained from GenBank and a further five *E. pilularis* sequences were obtained by cloning and sequencing as described in Sexton *et al.* (2010) (A list of the Genbank accession numbers and source organism is given in available from author upon request.

Once an approximate alignment was obtained, a conserved 636 bp fragment of the CCR gene was used to estimate the mean between group divergences amongst sections, subgenera or genera of eucalypts. The fragment represents conserved regions including Exon 4 and part of Intron 5 and Exon 5. Sequences were grouped at the genus, subgenus or section level according to the taxonomy of BROOKER (2000) as shown in *Supplemental Material 2*. The mean between-groups p-distance was estimated using the "compute between group means" and the "pairwise deletion" options in the Distance module of Mega v4.

Results

Transfer assessment – Stage 1

From the set of 71 loci tested for transfer, nine were not considered as candidates for transfer because they failed to amplify on the *E. grandis* control samples under the PCR conditions used in this study (*Table 1*; Rating zero). Ninety-seven percent (60) of the remaining 62 loci (*Table 1*; Rating 1 or higher) amplified in at least one of the six *E. pilularis*. The EPILMYB2 microsatellite marker also successfully amplified in both *E. grandis* and *E. pilularis* samples.

Of the 60 transferred loci, 16 could not be reliably scored due to a complex of peaks as a result of either multi-locus or non-specific amplification and were eliminated from further analysis (*Table 1*; Rating 3). A further three loci were monomorphic in *E. pilularis* (*Table 1*; Rating 2). Thirty-one loci were given the highest rating (5 or 6) for automated scoring suitability (*Table 2*), indicating that they were polymorphic, single-locus, and possessed electropherogram peak characteristics amenable to automated scoring. The remaining 10 loci were polymorphic but rated 4 indicating their characteristics were less ideal for automated analysis. The mean unbiased expected heterozygosity for these 41 markers (Rankings 4-6) was (mean (SD)) 0.75 (0.23). The EPILMYB2 marker developed for this study from *E. pilularis* sequence was also polymorphic and possessed characteristics suitable for automated scoring, with an unbiased expected heterozygosity of 0.91.

Transfer assessment – Stage 2

Twenty four loci, including the EPILMYB2 marker, were selected from the pool of markers with a rating of 5 or 6 on the basis of the compatibility of their T_a and PCR product size for grouping into sets of loci suitable for multi-plex PCR. These markers were re-assessed for amenability to automated scoring and polymorphism. Two loci, EMBRA 204 and EMBRA 182 were eliminated from further analysis as they could not to be scored and assessed for diversity due to weak signal levels under the pooled analysis of PCR products used here (*Table 3*). The genetic diversity of remaining transferred loci was not different (n = 21; av. $UH_e = 0.81 \pm 0.19$; one-sided t test unequal variances p-value = 0.22; nb EPILMYB2, EMBRA 204 and 182 excluded; *Table 3*) from the entire set of transferred and polymorphic loci (n = 41; see above).

Analysis of parental and offspring genotypes for the controlled pollinated pedigree revealed that no non-parental alleles were observed in the F_1. However, null alleles probably did occur in EMBRA 070 (data not shown). Most loci (18 out the 24) did not exhibit the

Table 4. – Pairwise distances between *Angophora*, *Corymbia* and *Eucalyptus* based on sequence diversity within the CCR gene (636bp).

	Angophora	*Corymbia*					*Eucalyptus*			
	A	*CB*	*CR*	*CP*	*CC*	*CO*	*EP*	*EA*	*SE*	*SM*
A										
CB	0.05									
CR	0.04	0.03								
CP	0.05	0.03	0.02							
CC	0.04	0.01	0.01	0.01						
CO	0.04	0.04	0.02	0.02	0.01					
EP	0.15	0.12	0.14	0.14	0.14	0.14				
EA	0.13	0.12	0.12	0.12	0.12	0.14	0.00			
SE	0.12	0.12	0.12	0.11	0.11	0.13	0.05	0.05		
SM	0.13	0.12	0.12	0.11	0.11	0.13	0.05	0.04	0.02	
SL	0.13	0.12	0.12	0.11	0.11	0.13	0.05	0.04	0.02	0.01

Taxa codes: A = Genus *Angophora*; CP, CB, CR,CC, CO are sections Politaria, Blakearia, Rufaria, Cadagaria and Ocharia of the genus Corymbia; *SE*, *SM* and *SL* are the sections, *Exertaria*, *Maidenaria* and *Latoangulatae* of the genus *Eucalyptus*, subgenus *Symphyomyrtus*; EP and EA are sections *Pseudophloius* and *Aromatica* of genus *Eucalyptus* subgenus *Eucalyptus*.

Table 5. – Rates of transfer of microsatellite markers at different taxonomic levels within eucalypts.

Focal taxa	*Non-focal taxa*	*No. of loci*	*Within sub-genera (%)*	*Across subgenera (%)*	*Across genera (%)*	*Reference*
E. nitens	*C. maculata 2 Mono. 4 Symph.*	4	100	50	nil	(BYRNE *et al.*, 1996)
E. grandis	*E. dunnii, E. cloeziana E. pilularis C. citriodora*	100	80-100	55-59	20	(KIRST *et al.*, 1997)
Corymbia citriodora spp variegata	*C. henryi, C. watsoniana, C. intermedia E. grandis E. globulus E. nitens E. pilularis E. acmenoides E. cloeziana*	14	100	57-79	21-53	(JONES *et al.*, 2001)[1]
E. grandis	*Corymbia* ssp.	73			90	(SHEPHERD *et al.*, 2006)
E. sieberi	*E. nitens*	8		75		(GLAUBITZ *et al.*, 2001)
E. globulus	*6 Sympho. 4 Mono. 2 Corymbia spp. 2 Angophora*	13	88	69	53-69	(STEANE *et al.*, 2001)[2]
E. grandis	*E. pilularis*	71		96		This study

[1] Transfer rates may be inflated by non-specific PCR products.
[2] Transfer rates exclude non-specific PCR products.

allele sizes consistent with the size of their nucleotide repeat motif. For example, loci having di-nucleotide repeats were found to have both odd and even-sized alleles (*Table 3*).

An independent estimate of the divergence amongst eucalypt groups at a fragment of the CCR gene.

Sequence variation in an alignment of CCR gene was used to generate estimates of divergence between members of the *Symphyomyrtus* and *Eucalyptus* subgenera as a comparison to divergence inferred by the transfer rate for microsatellite markers (*Table 4*). A total of 636 bp from the CCR gene, consisting of 172 bp of intronic and 464 bp exonic sequence from Exons 4 and 5 and Intron 4 (alignment not shown) was aligned for 55 eucalypts (See *Supplemental Material 1* for list of species and Genbank accession numbers). The percentage divergence between the source section for microsatellite marker origin (i.e. Section *Latoangulatae* in the Subgenus *Symphyomyrtus*; estimate based on 12 individuals from 8 species) and target section (Section *Pseudophloius* from the Subgenus *Eucalyptus*; estimate based on 5 *E. pilularis*) averaged 4% and was congruent with the level of divergence between another subgenus *Eucalyptus* representative, *E. amygdalina* (n = 1) and Section *Latoangulatae* representatives (*Table 4*). The level of divergence evident between subgenera therefore, was around two fold that of the largest divergence amongst members from different sections of the subgenus *Symphyomyrtus* (i.e. between SE and SM, or SE and SL; ie. both were 2%; *Table 4*). We also note that the divergence between *Eucalyptus* sections and *Corymbia* sections or representatives of the *Angophora* genus, ranged between 0.11 and 0.14%, which was about 2–3 fold the magnitude of divergence evident among subgenera of *Eucalyptus* (0.04–0.05). Also noteworthy was the consistently lower inter-generic divergences evident between representatives of *Angopho-*

ra and *Corymbia* genera (i.e. 0.04 to 0.05), compared with the *Eucalyptus* to *Angophora* or *Corymbia* contasts, consistent with the more recent view of sister taxa relationship of *Angophora* and *Corymbia* (*Table 4*).

Discussion

Assessing and reporting microsatellite transfer rates provides information on the evolution of microsatellite loci and should lead to methods that improve transfer rates (PRIMMER *et al.,* 2005). We found that transfer of microsatellite loci across subgenera of *Eucalyptus* can be highly efficient, with 60 out 62 loci that transferred to the laboratory (i.e. amplified on the source taxon ie. *E. grandis*), transferred across subgenera, and all but 3 exhibited polymorphism in the target taxon.

The transfer rate for the present study tended to be high compared to previously published studies of subgeneric or wider transfer of microsatellite marker in eucalypts which have generally been moderate (i.e. 50–75%) (*Table 5*). A key exception was a study of intergeneric transfer between *Eucalyptus* and *Corymbia* where a transfer rate of 90% was recorded (SHEPHERD *et al.,* 2006). These two studies were unique among the tabulated examples in that transfer was assessed in a laboratory other than the one where the markers were developed. Consequently, prior to estimation of transfer efficiencies, loci had to be transferred to the laboratory and amplify on control samples of the source taxon under conditions where there were both deliberate and unintended (i.e. differences between PCR machines, enzymes or other reagents) modifications to the PCR parameters.

In the present study, nine loci were eliminated for consideration in the transfer study because they did not amplify on the source taxon under the imposed PCR conditions. Hence the process of transferring loci among laboratories may have the unintended consequences of filtering for loci with wider PCR optima. In doing so, this process of filtering is further likely to identify loci that are more robust to less-than-perfect primer-template matching, thus leading to inadvertent selection of loci more likely to transfer across species. This process of ascertainment bias, may have helped identify loci that will transfer more frequently across species, and have contributed to the high rates of transfer recorded for the two studies in question.

The adoption of the technique of touchdown PCR may have been a second factor augmenting high realised transfer rates in the present study relative to some earlier studies. Touchdown PCR is a technique invented to circumvent the need to empirically optimise a fixed annealing temperature (T_a) for PCR, because of the difficulties of estimating melting temperatures for templates, or because the template sequence is unknown, hence it a technique particularly valuable for cross-species transfer of loci (ROUX, 1994). Optimal T_a may not have been tested in the two earlier transfer studies that used fixed T_a (JONES *et al.,* 2001; STEANE *et al.,* 2001) as neither report empirical testing for a T_a optimum. Higher rates for studies adopting touchdown PCR for cross-species amplification, where the aim is to commence cycling at a T_a above the T_a optimum and undergo stepwise decreases as cycling progresses, is consistent with studies of increased success rates for microsatellite markers transfer in birds when fixed annealing temperatures are lowered (PRIMMER *et al.,* 1995).

In addition to the technical factors discussed, high transfer rates would not be possible were the background biological factors also not suitable. The success of cross-species transfer of microsatellite markers is related to evolutionary distance of the transfer which affects both the divergence in primer site sequence, and insertions in the intervening sequence, and therefore the success of PCR (HEDGECOCK *et al.,* 2004; PRIMMER *et al.,* 2005).

Early predictions of high rates of microsatellite marker transfer among eucalypts based on experience with relatively small numbers of microsatellite markers and sequence conservation evident at other genetic loci have largely been borne out (BYRNE *et al.,* 1996; GLAUBITZ *et al.,* 2003; KIRST *et al.,* 1997). And, as the present study shows, despite the difficulties with using SSR transfer rates as a measure of evolutionary divergence (due to ascertainment bias for example), high transfer rates are nonetheless congruent with the accumulating evidence of high sequence conservation in *Eucalyptus*.

There remains relatively few published phylogenetic studies spanning subgenera of *Eucalyptus* or the eucalypts more broadly (*Eucalyptus*, *Corymbia* and *Angophora*) (Recently reviewed in GRATTAPAGLIA *et al.* (2012)). The study based on internal transcribed spacer (ITS) sequence remains among the most comprehensive, and indicate sequence divergence between subgenera in eucalypts tends to the lower end of that found in sections or genera of other angiosperms (STEANE *et al.,* 1999; See *Table 2* of BALDWIN et al. (1995)). Estimates of percentage pairwise divergence between two members of *Eucalyptus* subgenera ranged from 1.2 to 3% and were not much different to those within each subgenus (1.9% *Symphyomyrtus*; 1.0% *Monocalyptus*).

Our assessment of subgenus divergence using the CCR gene revealed remarkably similar and low sequence divergence among *Eucalyptus* subgenera. The percentage divergence in CCR between the source group (Section *Latoangulatae* n = 12 individuals from 8 species) and target (Section *Pseudophloius* n = 5 *E. pilularis*), averaged 4%, equal to that between another subgenus *Eucalyptus* representative, *E. amygdalina* (n = 1) and Section *Latoangulatae* representatives.

Our analysis of the CCR fragment also provided estimates of inter-generic divergence that revealed divergence between *Eucalyptus* and *Angophora* or *Corymbia* was 2 to 3 fold that among the *Eucalyptus* subgenera. This again augers well for inter-generic SSR transfer, as was previously reported, but clearly the relationship between sequence divergence and transfer rates is not linear and many other factors influence this relationship. As an aside, we note also the closer affinity between the *Corymbia* and *Angophora* genera consistent with more recent views of their taxonomic status (HILL and JOHNSON, 1995; LADIGES and UDOVICIC, 2000), and

favourable prospects for transfer of SSR loci developed in *Corymbia* to its sister genus (SHEPHERD *et al.,* 2006).

The present study reports a set of 22 well-characterised microsatellite markers for use in *E. pilularis*. These loci are likely to be robust to fluctuations in PCR conditions and amenable to use in other species as well, a function of the inadvertent ascertainment bias that occurs when loci are transferred among laboratories that may favour loci with broad PCR optima. Although issues of ascertainment bias muddle estimates of evolutionary divergence, high microsatellite transfer rates for eucalypts are in keeping with the accumulating evidence of high sequence level conservation in the group.

Funding

This study was funded by the Cooperative Research Centre for Forestry and Forests NSW.

Acknowledgements

The author thanks Forests NSW staff and contractors, in particular, S. BOYTON, R. ALLEN, R. SHIELS, H. SMITH, G. HEAGNEY, C. MORAN, W. VARLEY, D. KLIENIG and J. ROSS for help in foliage collections and acknowledge the assistance of the DNA Bank, Southern Cross Plant Genomics, and K. NORTHCOTT for the assistance with laboratory work, T. SEXTON and L. MCMANUS for development of the EPILMYB2 locus, and C. Harwood for comments on the manuscript.

References

AUCKLAND, L. D., T. BUI, Y. ZHOU, M. SHEPHERD and C. G. WILLIAMS (2002): Conifer Microsatellite Handbook. Corporate Press, Raleigh NC.

BALDWIN, B. G., M. J. SANDERSON, J. M. PORTER, M. F. WOJCIECHOWSKI and C. S. CAMPBELL *et al.* (1995): The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. Ann Missouri Botanical Gardens **82**: 247–277.

BARBARA, T., C. PALMA-SILVA, G. M. PAGGI, F. BERED and M. F. FAY **et al.** (2007): Cross-species transfer of nuclear microsatellite markers: potential and limitations. Molecular Ecology **16**: 3759–3767.

BRONDANI, R. P. V., C. BRONDANI and D. GRATTAPAGLIA (2002): Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. Molecular Genetics & Genomics **267**: 338–347.

BRONDANI, R. P. V., C. BRONDANI, R. TARCHINI and D. GRATTAPAGLIA (1998): Development, characterisation and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. Theoretical and Applied Genetics **97**: 816–827.

BRONDANI, R. P. V., E. R. WILLIAMS, C. BRONDANI and D. GRATTAPAGLIA (2006): A microsatellite-based consensus linkage map for species of *Eucalyptus* and a novel set of 230 microsatellite markers for the genus. BMC Plant Biology **6**: 1–16.

BROOKER, M. I. H. (2000): A new classification of the genus *Eucalyptus* L'Her. (Myrtaceae). Australian Systematic Botany **13**: 79–148.

BYRNE, M., M. I. MARQUEZ-GARCIA, T. UREN, D. S. SMITH and G. F. MORAN (1996): Conservation and genetic diversity of microsatellite loci in the genus *Eucalyptus*. Australian Journal of Botany **44**: 331–341.

ELDRIDGE, K., J. DAVIDSON, C. HARWOOD and G. VAN WYK (1994): Eucalypt domestication and breeding. Oxford University Press, Oxford.

GLAUBITZ, J., EMEBIRI and G. MORAN (2001): Dinucleotide microsatellites in *Eucalyptus sieberi*: Inheritance, diversity and improved scoring of single base differences. Genome **44**: 1041–1045.

GLAUBITZ, J. C., H. X. WU and G. F. MORAN (2003): Impacts of silviculture on genetic diversity in the native forest species *Eucalyptus sieberi*. Conservation Genetics **4**: 275–287.

GRATTAPAGLIA, D. and M. KIRST (2008): *Eucalyptus* applied genomics: from gene sequences to breeding tools. New Phytologist **179**: 911–929.

GRATTAPAGLIA, D., R. E. VAILLANCOURT, M. SHEPHERD, B. THUMMA and W. J. FOLEY *et al.* (2012): Progress in Myrtaceae genomics; *Eucalyptus* as the pivotal genus. Tree Genetics and Genomes **8**: 463–508.

HEDGECOCK, D., G. LI, S. HUBERT, K. BUCKLIN and V. RIBES (2004): Widespread null alleles and poor cross-species amplification of microsatellite DNA loci cloned from the Pacific oyster, *Crassostrea gigas*. Journal of Shellfish Research **23**: 379–385.

HENSON, M. and H. J. SMITH (2007): Achievements in forest tree genetic improvement in Australia and New Zealand. 1. *Eucalyptus pilularis* Smith tree improvement in Australia. Australian Forestry **70**: 4–10.

HILL, K. D. and L. A. S. JOHNSON (1995): Systematic studies in the eucalypts 7. A revision of the bloodwoods, genus *Corymbia* (Myrtaceae). Telopea **6**: 185–504.

JONES, M., R. STOKOE, M. CROSS, L. SCOTT and T. MAGUIRE *et al.* (2001): Isolation of microsatellite loci from spotted gum (*Corymbia variegata*), and cross-species amplification in *Corymbia* and *Eucalyptus*. Molecular Ecology Notes **1**: 276–278.

KIRST, M., R. P. V. BRONDANI, C. BRONDANI and D. GRATTAPAGLIA (1997): Screening of designed primer pairs for recovery of microsatellite markers and their transferability among species of *Eucalyptus*, pp. 167–170. *In:* Proceedings of the IUFRO Conference on Silviculture and Improvement of Eucalypts, Salvador, Brazil.

LADIGES, P. Y. and F. UDOVICIC (2000): Comment on a new classification of the eucalypts. Australian Systematic Botany **13**: 149–152.

NEI, M. (1978): Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89**: 583–590.

PEAKALL, R., S. GILMORE, W. KEYS, M. MORGANTE and A. RAFALSKI (1998): Cross-species amplification of soybean (*Glycine* max) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. Molecular Biology and Evolution **15**: 1275–1287.

PEAKALL, R. and P. E. SMOUSE (2006): GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes **6**: 288–295.

POKE, F. S., R. E. VAILLANCOURT, B. M. POTTS and J. B. RIED (2005): Genomic research in *Eucalyptus*. Genetica **125**: 79–101.

PRIMMER, C. R., A. P. MOLLER and H. ELLEGREN (1995): Resolving Genetic Relationships With Microsatellite Markers – a Parentage Testing System For the Swallow Hirundo Rustica. Molecular Ecology **4**: 493–498.

PRIMMER, C. R., J. N. PAINTER, M. T. KOSHINEN, J. U. PALO and J. MERILA (2005): Factors affecting avain cross-species microsatellite amplification. Journal of Avian Biology **36**: 348–360.

ROUX, K. H. (1994): Using mismatched primer-template pairs in touchdown PCR. BioTechniques 16: 812-814.

SCHLOTTERER, C. (2001): Genealogical inference of closely related species based on microsatellites. Genetic Research, Cambridge **78**: 209–212.

SEXTON, T., R. J. HENRY, L. J. MCMANUS, S. BOWEN and M. SHEPHERD (2010): Capture of assay template by multiplex PCR of long amplicons for genotyping SNPs and InDels with MALDI-TOF mass spectrometry. Molecular Breeding **25**: 471–480.

SHEPHERD, M., S. KASEM, D. LEE and R. HENRY (2006): Construction of microsatellite linkage maps for *Corymbia*. Silvae Genetica **55**: 228–238.

SHEPHERD, M. and C. A. RAYMOND (2010): Species delineation and gene flow in the Blackbutts (Genus *Eucalyptus* Subgenus *Eucalyptus* Section *Pseudophloius*). Conservation Genetics **11**: 1965–1978.

SHEPHERD, M., T. R. SEXTON, D. THOMAS, M. HENSON and R. J. HENRY (2010): Geographical and historical determinants of microsatellite variation in *Eucalyptus pilularis*. Canadian Journal of Forest Research **40**: 1051–1063.

STEANE, D. A., G. E. MCKINNON, R. E. VAILLANCOURT and B. M. POTTS (1999): ITS sequence data resolve higher level relationships among the eucalypts. Molecular Phylogenetics and Evolution **12**: 215–223.

STEANE, D. A., R. E. VAILLANCOURT, J. RUSSELL, W. POWELL, D. MARSHALL *et al.* (2001): Development and characterisation of microsatellite loci in *Eucalyptus globulus* (Myrtaceae). Silvae Genetica **50**: 89–91.

TAMURA, K., J. DUDLEY, M. NEI and S. KUMAR (2007): MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution **24**: 1596–1599.